(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 February 2001 (15.02.2001)

PCT

(10) International Publication Number
WO 01/10903 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/US00/21878

(22) International Filing Date: 9 August 2000 (09.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/147,986	9 August 1999 (09.08.1999)	US
60/160,807	21 October 1999 (21.10.1999)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/147,986 (CIP)
Filed on	9 August 1999 (09.08.1999)
US	60/160,807 (CIP)
Filed on	21 October 1999 (21.10.1999)

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **LAL, Preeti** [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). **TANG, Y., Tom** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **AZIMZAI, Yalda** [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). **LU, Dyung, Aina, M.** [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). **YANG, Junming** [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/10903 A2

(54) Title: PROTEASES AND PROTEASE INHIBITORS

(57) Abstract: The invention provides human proteases and protease inhibitors (PPIM) and polynucleotides which identify and encode PPIM. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of PPIM.

to distinguish between SEQ ID NO:28-54 and related polynucleotide sequences. The polypeptides encoded by the selected fragments of SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, and SEQ ID NO:54 are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express PPIM as a fraction of total tissues expressing PPIM. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing PPIM as a fraction of total tissues expressing PPIM. Column 5 lists the vectors used to subclone each cDNA library. Of particular note is the expression of SEQ ID NO:28 in gastrointestinal tissue. Of particular note is the tissue-specific expression of SEQ ID NO:51. Over 83% of the tissues expressing SEQ ID NO:51 are derived from gastrointestinal tissue, particularly the liver.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding PPIM were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:30 maps to chromosome 9 within the interval from 78.4 to 90.6 centiMorgans. This interval also contains a gene associated with cell proliferation.

SEQ ID NO:37 maps to chromosome 12 within the interval from 116.6 to 118.9 centiMorgans. This interval also contains a gene associated with a neurological disorder.

SEQ ID NO:47 maps to chromosome 4 within the interval from 99.2 to 105.2 centiMorgans. This interval also contains a gene associated with cardiovascular disease.

The invention also encompasses PPIM variants. A preferred PPIM variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the PPIM amino acid sequence, and which contains at least one functional or structural characteristic of PPIM.

The invention also encompasses polynucleotides which encode PPIM. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:28-54, which encodes PPIM. The polynucleotide sequences of SEQ ID NO:28-54, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding PPIM. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
21	48	3092341	BRSTNOT19	3092341H1 (BRSTNOT19), 3092341T6 (BRSTNOT19)
22	49	3658034	ENDPNOT02	2623516R6 (KERANOT02), 3658034F6 (ENDPNOT02), 3658034H1 (ENDPNOT02), 3658034T6 (ENDPNOT02), 5216522H1 (BRSTNOT35), 5590053H1 (ENDINOT02)
23	50	3883861	UTRSNOT05	858111H1 (NGANNOT01), 858233H1 (NGANNOT01), 1364808R1 (SCORNON02), 1861181F6 (PROSNOT19), 1906985T6 (OVARNOT07), 2687868H1 (LUNGNOT23), 2687868X366D1 (LUNGNOT23), 2721116X369D1 (LUNGTUT10), 3883861H1 (UTRSNOT05), 5217169H1 (BRSTNOT35)
24	51	4993873	LIVRTUT11	4987943H1 (LIVRTUT10), 4993873H1 (LIVRTUT11), SCEA01665V1, SCEA00232V1, SXBC01625V1, SXBC01802V1, SCSA03627V1
25	52	5208004	BRAFNOT02	4696870F6 (BRALNOT01), 5208004H1 (BRAFNOT02)
26	53	5267783	BRAFDIT02	220636R1 (STOMNOT01), 679457R6 (UTRSNOT02), 1330537F6 (PANCNOT07), 1808720F6 (PROSTUT12), 1969475H1 (BRSTNOT04), 2697426F6 (UTRSNOT12), 2991180H1 (KIDNFET02), 3532849H1 (KIDNNOT25), 4992376F6 (LIVRTUT11), 5004695F6 (PROSTUT21), 5267783H1 (BRAFDIT02)
27	54	5583922	FIBAUNT01	726878R1 (SYNOOAT01), 956818X11 (KIDNNOT05), 1658964X12 (URETTUT01), 1658964X13 (URETTUT01), 2544879F6 (UTRSNOT11), 3748858H1 (UTRSNOT18), 4761921H1 (PLACNOT05), 5043801H1 (PLACFER01)

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
48	271-315	Gastrointestinal (0.278) Reproductive (0.278) Cardiovascular (0.111) Hematopoietic/Immune (0.111) Nervous (0.111)	Cancer (0.444) Inflammation/Trauma (0.555) Cell Proliferation(0.167)	pINCY
49	217-261	Hematopoietic/Immune (0.364) Reproductive (0.273)	Cell Proliferation(0.364) Inflammation/Trauma (0.364) Cancer (0.182)	pINCY
50	164-208	Reproductive (0.333) Nervous (0.222) Gastrointestinal (0.167)	Cancer (0.611) Inflammation/Trauma (0.223)	pINCY
51	388-432	Gastrointestinal (0.833) Reproductive (0.166)	Cancer (0.666) Cell Proliferation(0.166)	pINCY
52	218-262	Nervous (0.750) Hematopoietic/Immune (0.250)	Inflammation/Trauma (0.500) Neurological (0.250)	pINCY
53	325-369	Reproductive (0.289) Nervous (0.253) Gastrointestinal (0.120)	Cancer (0.410) Inflammation/Trauma (0.386) Cell Proliferation(0.145)	pINCY
54	165-209	Reproductive (0.352) Urologic (0.185) Developmental (0.130)	Cancer (0.630) Cell Proliferation(0.167) Inflammation/Trauma (0.204)	pINCY

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
47	THP1AZS08	This subtracted THP-1 promonocyte cell line library was constructed using 5.76 million clones from a 5-aza-2'-deoxycytidine (AZ) treated THP-1 cell library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. The hybridization probe for subtraction was derived from a similarly constructed library, made from RNA isolated from untreated THP-1 cells. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (1991) Nucleic Acids Res. 19:1954; and Bonaldo et al. (1996) Genome Research 6:791. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia.
48	BRSTNOT19	Library was constructed using RNA isolated from breast tissue removed from a 67-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated residual invasive lobular carcinoma. The focus of residual invasive carcinoma was positive for both estrogen and progesterone. Patient history included depressive disorder and benign large bowel neoplasm. Family history included cerebrovascular disease, benign hypertension, congestive heart failure, and lung cancer.
49	ENDPNOT02	Library was constructed using RNA isolated from pulmonary artery endothelial cells removed from a 10-year-old Caucasian male. The cells were treated with TNF alpha and IL-1 beta 10ng/ml each for 20 hours.
50	UTRSNOT05	Library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, and SEQ ID NO:54.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, and SEQ ID NO:54,

b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a

polynucleotide sequence selected from the group consisting of SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, and SEQ ID NO:54,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4,

gttaatattc ctgctgtgag aaactttaaa gtttcaaata cccaagatgc ttcagtgtcc 660
atagtggatt actatgagcc aaggagacag gcggtgagaa gttacaactc tgaagtgaag 720
ctgtcctcct gtgacctttg cagtgatgtc cagggctgcc gtccttgtga ggatggagct 780
tcaggctccc atcatcactc ttcagtcatt tttatttttct gtttcaagct tctgtacttt 840
atggaacttt ggctgtgatt tatttttaaa ggactctgtg taacactaac atttccagta 900
gtcacatgtg attgtttttgt tttcgtagaa gaatactgct tctattttga aaaaagagtt 960
ttttttcttt ctatggggtt gcagggatgg tgtacaacag gtcctagcat gtatagctgc 1020
atagatttct tcacctgatc tttgtgtgga agatcagaat gaatgcagtt gtgtgtctat 1080
attttcccct ctcaaaatct tttagaattt ttttggaggt gtttgttttc tccagaataa 1140
aggtattact ttagaatagg tattctcctc attttgtgaa agaaatgaac ctagattctt 1200
aagcattatt acacatccat gtttgcttaa agatggattt ccctgggaat gggagaaaac 1260
agccagcagg aggagcttca tctgttccct tcccacctcc 1300

<210> 50
<211> 2241
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3883861CB1

<400> 50
gctccccgag ccgggctgca ccggaggcgg cgagatggtc gcgcgcgtcg gcctcctgct 60
gcgcgccctg cagctgctac tgtggggcca cctggacgcc cagcccgcgg agcgcggagg 120
ccaggagctg cgcaaggagg cggaggcatt cctagagaag tacggatacc tcaatgaaca 180
ggtccccaaa gctcccacct ccactcgatt cagcgatgcc atcagagcgt ttcagtgggt 240
gtcccagcta cctgtcagcg gcgtgttgga ccgcgccacc ctgcgccaga tgactcgtcc 300
ccgctgcggg gttacagata ccaacagtta tgcggcctgg gctgagagga tcagtgactt 360
gtttgctaga caccggacca aaatgaggcg taagaaacgc tttgcaaagc aaggtaacaa 420
atggtacaag cagcacctct cctaccgcct ggtgaactgg cctgagcatc tgccggagcc 480
ggcagttcgg ggcgccgtgc gcgccgcctt ccagttgtgg agcaacgtct cagcgctgga 540
gttctgggag gccccagcca caggccccgc tgacatccgg ctcaccttct tccaagggga 600
ccacaacgat gggctgggca atgcctttga tggcccaggg ggcgccctgg cgcacgcctt 660
cctgccccgc cgcggcgaag cgcacttcga ccaagatgag cgctggtccc tgagccgccg 720
ccgcgggcgc aacctgttcg tggtgctggc gcacgagatc ggtcacacgc ttggcctcac 780
ccactcgccc gcgccgcgcg cgctcatggc gccctactac aagaggctgg gccgcgacgc 840
gctgctcagc tgggacgacg tgctggccgt gcagagcctg tatgggaagc ccctaggggg 900
ctcagtggcc gtccagctcc caggaaagct gttcactgac tttgagacct gggactccta 960
cagccccaa ggaaggcgcc ctgaaacgca gggccctaaa tactgccact cttccttcga 1020
tgccatcact gtagacaggc aacagcaact gtacattttt aaagggagcc atttctggga 1080
ggtggcagct gatggcaacg tctcagagcc ccgtccactg caggaaagat gggtcgggct 1140
gccccccaac attgaggctg cggcagtgtc attgaatgat ggagatttct acttcttcaa 1200
agggggtcga tgctggaggt tccggggccc caagccagtg tggggtctcc cacagctgtg 1260
ccgggcaggg ggcctgcccc gccatcctga cgccgccctc ttcttccctc ctctgcgccg 1320
cctcatcctc ttcaagggtg cccgctacta cgtgctggcc cgagggggac tgcaagtgga 1380
gccctactac ccccgaagtc tgcaggactg gggaggcatc cctgaggagg tcagcggcgc 1440
cctgccgagg cccgatggct ccatcatctt cttccgagat gaccgctact ggcgcctcga 1500
ccaggccaaa ctgcaggcaa ccacctcggg ccgctgggcc accgagctgc cctggatggg 1560
ctgctggcat gccaactcgg ggagcgccct gttctgaagg cacctcctca cctcagaaac 1620
tggtggtgct ctcagggcaa aatcatgttc cccacccccg gggcagaacc cctcttagaa 1680
gcctctgagt ccctctgcag aagaccgggc agcaaagcct ccatctggaa gtctgtctgc 1740
ctttgttcct tgaagaatgc agcattgtct ttgtctgtcc ccaccacatg gaggtggggg 1800
tgggatcaat cttaggaaaa gcaaaaaagg gtcccagatc ccttggccct ttcctccgag 1860
gacttctatc ctccccaggc ctttgtttct tcggctaaag gtacagttcc tttcaagagg 1920
taacagcact gggatccaag caggggggatg aaaaactcag cagagaaatt cgagaccatt 1980
ttgcaagact gtgcccttct cctcaggacc ccctggctca gttcttgaaa aacggtgtca 2040
tatttagtca gaggccccac ccccaggaag catggatggg gatgaaggca caggcgtctc 2100
caacctcaga ggccctttgt ggggtcagga cacagagtgg gagggagact gatgcaggcc 2160
taccagtccc tggctttttg tctggggctg gaataaagag gtgccttcag ctggtgggcc 2220
gagaggcagg aaaaaaaaaa a 2241

<210> 51
<211> 1860
<212> DNA
<213> Homo sapiens